ON A FUNDAMENTAL THEOREM IN MULTIPLE COMPARISONS

by

E. L. Lehmann*

Juliet Popper Shaffer**

N00014-75-C-0444
NSF-MPS-73-08698
NR-042-036
Jan 76
9p.

A recent paper by Einot and Gabriel [1] restates a basic result for multiple range tests, which had earlier been given by Tukey [6] and Spjøtvoll [5]. It is pointed out that this result requires for its validity an assumption which is not explicitly provided by any of these earlier versions. There is also some discussion of a related result.

JAN 1976

**E. L. Lehmann is professor at Department of Statistics, University of California, Berkeley, Calif. 94720. Juliet Popper Shaffer is professor at Department of Psychology, University of Kansas, Lawrence, Kans. 66044 and at present visiting at Department of Mathematics, University of California, Davis, Calif. 95616.
*This paper was prepared with support of U.S. Office of Naval Research Contract No. N00014-75-C-0444/NR042-036. Also this research was prepared with the support of National Science Foundation, Grant No. MPS 73-08698.

Let $X_i (i=1,\ldots,s)$ be independently distributed with distribution

$$P(X_i \leq x) = F(x-\theta_i) , \tag{1}$$

and let Y_i denote the i^{th} smallest of the X's. We shall be concerned with the problem of grouping the θ's by means of a multiple range test, defined in terms of critical values $C_2,\ldots,C_s$ as follows. As a first step the range $R_s=Y_s-Y_1$ is compared with C_s: If $R_s<C_s$, the θ's are declared indistinguishable and the procedure terminates; if $R_s \geq C_s$, the θ's corresponding to Y_1 and Y_s are declared to differ, and the two $(s-1)$-ranges

$$R_{s-1,1} = Y_{s-1}-Y_1 \quad \text{and} \quad R_{s,2} = Y_s-Y_2$$

are compared with C_{s-1}. If both R's are less than C_{s-1} the two sets of θ's corresponding to $(Y_1,Y_2,\ldots,Y_{s-1})$ and $(Y_2,\ldots,Y_s)$ are declared indistinguishable and the procedure terminates. Otherwise the two θ's corresponding to Y_1,Y_{s-1} and/or the two θ's corresponding to Y_2,Y_s are declared to differ, and the three $(s-2)$-ranges are compared with C_{s-2}; and so on. After a set of means has been declared indistinguishable all of its subsets are also considered indistinguishable without further test.

To complete specification of the procedure, it is necessary to decide on the critical values $C_2,\ldots,C_s$. This choice is typically made in terms of the probabilities

$$\alpha_k = P[R_k \geq C_k] \tag{2}$$

where R_k denotes the range of k independent X's with common location parameter θ.

A quantity of central interest to many investigators of multiple range tests is the probability of declaring at least one pair of θ's to differ when they are in fact equal, say

$$\alpha_0 = \sup \alpha(\theta_1, \ldots, \theta_s) \tag{3}$$

where

$$\alpha(\theta_1, \ldots, \theta_s) = P_{\theta_1, \ldots, \theta_s}[\text{at least one false significance statement}]. \tag{4}$$

The fundamental theorem referred to in the title relates α_0 to $\alpha_2, \ldots, \alpha_k$. Suppose that the θ's fall into t groups of equal values of lengths $v_1, \ldots, v_t$; without loss of generality let the θ's be numbered so that

$$\theta_1 = \cdots = \theta_{v_1};\ \theta_{v_1+1} = \cdots = \theta_{v_1+v_2};\ \ldots \tag{5}$$

with the values in the different groups being distinct. Then the fundamental theorem states that

$$\sup \alpha(\theta_1, \ldots, \theta_s) = 1 - \prod_{i=1}^{t} (1 - \alpha_{v_i}) \tag{6}$$

where the sup is taken over all θ's which after reordering satisfy (5) and where $\alpha_1=0$.

This result appears to have been stated first, rather informally and with a sketch of a proof, in Chapter 3 of Tukey's unpublished book (1953) on multiple comparisons; it was discovered again by Spjøtvoll and plays a central role in his unpublished 1971 paper [5]; and now finally has appeared in the recent paper by Einot and Gabriel (1975).

The theorem in fact requires an additional assumption. A natural sufficient condition for its validity which the earlier authors may have assumed tacitly but which they did not state explicitly is

$$C_2 \le C_3 \le \dots \le C_s . \tag{7}$$

Actually, it is enough to assume that

$$C_2 \le C_3 \le \dots \le C_{\max v_i} \le C_j \quad \text{for all} \quad j \ge \max v_i . \tag{8}$$

That the theorem is not correct without some restriction on the C's, can be seen from the example $s=3$, $C_3=0$, $\theta_1=\theta_2 < \theta_3$ and $\theta_3 \to \theta_2$. Then the probability $\alpha(\theta_1,\theta_2,\theta_3)$ of declaring θ_1 to differ from θ_2 exceeds the probability that X_3 is between X_1 and X_2 (since in that case $|X_2-X_1|>C_3$) which is 1/3, regardless of the value of α_2 while the right hand side of (6) is α_2.

Condition (7) is of course trivially satisfied for the Newman-Keuls procedure defined by

$$\alpha_2 = \ldots = \alpha_s \tag{9}$$

and by Tukey's T-method defined by

$$C_2 = \ldots = C_s \,. \tag{10}$$

Inspection of Table B3 of Harter [2] suggests that it also holds for Duncan's procedure when F is normal. Whether in that case it is true for arbitrary F, we do not know.

Because of the importance of the fundamental theorem it may be worth making available a proof. The proof below is essentially that given by Spjøtvoll and sketched in the other papers.

Proof of (6), assuming (7).

When the θ's satisfy (5),

$$\alpha(\theta_1,\ldots,\theta_s) \leq P\left(\bigcup_{i=1}^{t} [R_i' \geq C_{v_i}]\right) \tag{11}$$

where R_i' denotes the range of the X's corresponding to the i^{th} group of θ's in (5). The right hand side of (11) is equal to

$$1 - P[R_i' \leq C_{v_i} \text{ for all } i] = 1 - \prod_{i=1}^{t} (1-\alpha_{v_i}) \tag{12}$$

which shows that the right hand side of (11) is an upper bound for $\alpha(\theta_1,\ldots,\theta_s)$.

Note that assumption (8) is needed to insure the validity of (11).

To prove that the right hand side of (6) is a sharp upper bound for $\alpha(\theta_1,\ldots,\theta_s)$ suppose that the difference between the values of successive groups in (5) is Δ. Then as $\Delta \to \infty$, the probability $\alpha(\theta_1,\ldots,\theta_s)$ tends to the right hand side of (11) and hence to (12) and this completes the proof.

Condition (7) plays a role in another result mentioned both by Tukey and in [1]. This is the fact that α_i is the maximum probability of declaring the set $\{\theta_1,\ldots,\theta_i\}$ to be nonhomogeneous when in fact $\theta_1 = \ldots = \theta_i$. That this result is not correct without some restriction is again shown by the example given after (8).

Let us now prove that it does hold under (7) or the weaker assumption

$$C_k \geq C_i \quad \text{for all} \quad k \geq i \; . \tag{13}$$

The probability of declaring the set $\{\theta_1,\ldots,\theta_i\}$ not to be homogeneous is

$$P[R_k \geq C_k \text{ for all ranges } R_k \text{ containing } X_1,\ldots,X_i] \tag{14}$$

$$\leq \Sigma\, p_k\, P[\max\,(X_1,\ldots,X_i) - \min\,(X_1,\ldots,X_i) \geq C_k]$$

where p_k is the probability that if

$$\min (X_1,\ldots,X_i) = Y_a \quad \text{and} \quad \max (X_1,\ldots,X_i) = Y_b$$

then $b-a = k-1$. If (13) holds, the right hand side of (14) is increased when C_k is replaced by C_i and then becomes α_i.

For applications, the special cases of the theorem when F is either normal or exponential are of particular interest. One may then wish to extend the theorem to the case of an unknown common scale parameter. What continues to be true is that

$$\alpha(\theta_1,\ldots,\theta_s) \leq 1 - \prod_{i=1}^{t} (1-\alpha_{v_i}) . \tag{15}$$

To see this, suppose that each X_i is replaced by X_i/S where S is an estimator of the scale parameter, which is assumed to be independent of the X's. Then the Studentized ranges $R_k^* = R_k/S$ are positively dependent (see for example, Lehmann [4, Ex. 1(ii)]), and the proof of (6) requires no changes until the last step, where equality in (12) must be replaced by an inequality: the left hand side is smaller than or equal to the right hand side.

Although the right hand side of (15) is no longer a sharp upper bound (and the upper bound may obtain for different configurations than in the proof of (6)), it seems likely that in the normal case, when σ is estimated with at least moderate degrees of freedom, the least upper bound is close to the right hand side (see e.g. Hartley [3]).

While conditions (7), (8) and (13) obviously continue to hold for the Newman-Keuls and for Tukey's T-method, they are no longer satisfied in the normal case with the probabilities

$$\alpha_k = 1 - (1-\alpha_2)^{k-1} \qquad (16)$$

proposed by Duncan. It appears from inspection of the Tables supplied for Duncan's procedure that he is in fact not advocating (16). Rather, he recommends the stated value of α_2 but lower values of α_k, $k>2$, as needed to get condition (7) to hold.

REFERENCES

[1] Einot, Israel, and Gabriel, K.R., "A study of the powers of several methods of multiple comparisons," Journal of the American Statistical Association, 70 (September 1975), 574-583.

[2] Harter, H.L., "Order statistics and their use in testing and estimation," Vol. 1, Washington: U.S. Government Printing Office, 1970.

[3] Hartley, H.O., "Some recent developments in analysis of variance," Communications in Pure and Applied Mathematics, 8 (February 1955), 47-72.

[4] Lehmann, E.L., "Some concepts of dependence," Annals of Mathematical Statistics, 37 (October 1966), 1137-1153.

[5] Spjøtvoll, Emil, "On the probability of at least one false rejection for various multiple testing techniques in the one-way layout in the analysis of variance," Statistical Research Report No. 6, Department of Mathematics, University of Oslo, 1971.

[6] Tukey, J.W., "The problem of multiple comparisons," Unpublished book, Princeton University, 1953.